ABSTRACT OF THE DISCLOSURE

Multimedia content containing moving pictures and audio is divided into multiple scenes and metadata is generated for each of the scenes obtained as a result of the division. It is possible to generate metadata containing scene structure information metadata that describes the hierarchical structure of the content in addition to scene section information and titles. Also, a name or an identifier of each descriptor contained in the metadata is described as hint information for manipulation of metadata composed of at least one descriptor describing semantic content, a structure, and characteristics of content.